(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 June 2002 (13.06.2002)

PCT

(10) International Publication Number
WO 02/46220 A2

(51) International Patent Classification[7]: C07K 14/47

(21) International Application Number: PCT/EP01/15412

(22) International Filing Date: 5 December 2001 (05.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/251,582 7 December 2000 (07.12.2000) US

(71) Applicants *(for all designated States except US)*: **AVENTIS PHARMA S.A.** [FR/FR]; 20 Avenue Raymond Aron, F-92160 Antony (FR). **THE REGENTS OF THE UNIVERSITY OF CALIFORNIA** [US/US]; 9500 Gilman Drive, La Jolla, CA 92093-0910 (US). **BENOIT, Patrick** [FR/FR]; 24 rue Jonquay, F-75014 Paris (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SCHWARTZ, Bertrand** [FR/FR]; 14 rue Georges Clémenceau, F-78350 Jouy En Josas (FR). **BRANELLEC, Didier** [FR/FR]; 5 rue Hector Berlioz, F-77170 Brie-Comte-Robert (FR). **CHIEN, Kenneth** [US/US]; 6232 Calle Vera Cruz, La Jolla, CA 92037 (US).

(74) Agents: **LECCA, Patricia** et al.; Aventis Pharma S.A., Patent Department, 20 Avenue Raymond Aron, F-92165 Antony Cedex (FR).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/46220 A2

(54) Title: SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING THEM AND USES THEREOF

(57) Abstract: The invention relates to novel promoter sequences derived from a portion upstream of the coding sequence of the gene for the CARP protein (Cardiac Ankyrin Repeat Protein), and which are capable of controlling the level and the specificity of expression of a transgene *in vivo* in cardiac muscle cells. The invention thus describes novel compositions, constructs, vectors and their uses *in vivo* for the transfer and expression of a nucleic acid *in vivo* in cardiac muscle cells. The subject of the present invention is also the use of the promoter sequences for generating transgenic animals which constitute models for studying certain cardiac pathologies.

SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING THEM AND USES THEREOF

The present invention relates to the field of biology. It relates in particular to the field of the targeting of the expression of genes, and more particularly the design and the development of a novel system for the specific expression of transgenes. The subject of the invention is in particular novel promoter sequences capable of controlling the level and the specificity of expression of a transgene *in vivo* in cardiac muscle cells. The invention thus describes novel compositions, constructs and vectors which make it possible to control and to direct the expression of a nucleic acid in cardiac muscle cells. The applications stemming from the present invention are numerous, for example in the experimental, clinical, therapeutic and diagnostic fields, and more particularly for the treatment and/or prevention of certain cardiac pathologies.

The control of the level and of the targeting of the expression of the transgenes is necessary for many applications. Thus, in gene therapy, the success of the therapy may require targeting of the protein synthesized from the transgene and thus make it possible to limit the spreading of the side effects. The construction of transgenic animals, and the study of the effects of a gene are all examples in which an appropriate control of the specificity of expression of a protein can be used and can provide improvements.

In this regard, many promoters have been tested for their capacity to direct a cardiospecific expression.

They are in particular the promoter of the gene encoding the cardiac myosin light chain (MLC-2) in rats (Henderson S.A. et al., *J Biol Chem*, **264** (1989) 18142-8; Lee K.J. et al., *J Biol Chem*, **126** (1992) 15875-85), cardiac α-actin in mice (Biben C. et al., *Dev Biol*, **173** (1996) 200-12), natriuretic factor (ANF) (Harris A.N. et al., *J Mol Cell Cardiol*, **29** (1997) 515-25), α- or β-myosin heavy chain (α- or β-MHC) (Colbert M.C. et al., *J Clin Invest*, **100** (1997) 1958-68), muscle creatine kinase (MCK) in rabbits (Vincent C.K. et al., *Mol Cell Biol*, **13** (1993) 567-74), or cardiac troponin T (US 5,266,488).

While these promoters are known to confer a degree of tissue specificity, it is also known that their levels of activity remain well below those of so-called strong promoters, generally by a factor of between 10 and 100, such that a therapeutic use cannot really be envisaged.

By way of example, Franz W.M. et al., (*Cardiovasc Res*, **35** (1997) 560-6) and Griscelli F. et al., (*C R Acad Sci III*, **320** (1997) 103-12) have shown that the levels of activity of the sequences upstream of the genes encoding rat α-MHC and MLC-2 in adenoviral constructs remain substantially lower than those of the RSV (Rous sarcoma virus) promoter, by a factor of about 10.

The present application therefore relates more precisely to a novel promoter sequence derived from the region upstream of the CARP (Cardiac Ankyrin Repeat Protein) gene. It is not only capable of directing a cardiospecific expression, but also exhibits a high level of expression *in vivo* comparable to that of a strong promoter such as the CMV (cytomegalovirus) promoter.

The CARP protein, which constitutes one of the first markers for differentiation of the cardiomyocytes acting downstream of the homeobox gene Nbx2.5 in the regulation of the expression of the MLC-2v gene, has been studied and the coding portion of its gene has been sequenced in mice (Zou Y. et al., *Development*, **24** (1997) 793-804), in rabbits (Aihara Y. et al., *Biochim Biophys Acta*, **28** (1999) 318-24), and in humans (Chu W. et al., *J Biol Chem*, **270** (1995) 10236-45).

Kuo H. et al. (*Development*, **126** (1999) 4223-34) have cloned a fragment of 10 Kb and sequenced a fragment of 2.5 Kb upstream of the coding sequence of the mouse CARP gene. Deletions in 5′ were made in the fragment showing that a region of 213 bp of the promoter between nucleotides -166 and +47, relative to the transcription position +1, was sufficient to confer cardiospecific expression *in vitro*, which suggested the presence, at the 5′ end, of an element for controlling the specificity of the promoter. Kuo et al. also generated transgenic mouse lines comprising a fragment of 2.5 Kb upstream of the CARP gene, showing specific expression of a transgene in the cardiac and skeletal muscle cells at the early stage of embryonic development, this expression then being inhibited during development.

Application WO 00/15821 describes a portion in 5′ upstream of the coding sequence of the mouse CARP gene, situated between nucleotides -2285 and +62, relative to transcription position +1. This sequence was evaluated in particular for its *in vivo* activity via adenoviral vectors. The levels of activity obtained remain however very low, such that it was found to be necessary, in order

to detect an activity *in vivo*, to isolate the promoter sequence between two inverted terminal repeats of an adeno-associated virus (AAV-ITR).

The Applicant focused on better characterizing the region in 5′ of the gene for the CARP protein. It was thus able to identify a novel sequence upstream of the CARP gene and demonstrate unexpected and advantageous properties of this novel sequence, in particular a significant improvement in the levels of activity *in vivo*.

The Applicant has indeed discovered, surprisingly, that while this newly identified sequence conferred no significant expression *in vitro*, it was on the contrary possible to obtain very good levels of activity *in vivo*, equivalent to those of so-called strong promoters, while preserving a high selectivity of expression in the cardiac tissue.

The subject of the present invention is therefore a polynucleotide comprising a portion upsteam of the coding sequence of the gene for the CARP protein, or of a portion hybridizing under highly stringent conditions with the said upstream sequence, the said polynucleotide being capable of inducing specific expression in the cardiac tissue of a transgene placed under its control.

The invention also relates to any polynucleotide of natural origin or which is obtained by chemical synthesis, exhibiting at least 93%, preferably at least 95%, identity with the sequence SEQ ID NO: 1. More preferably still, the polynucleotide according to the invention exhibits at least 98% identity with the sequence SEQ ID NO: 1.

The expression polynucleotide of natural origin is understood to mean a genomic DNA fragment obtained by

cleaving cellular DNA with the aid of a restriction enzyme.

The expression polynucleotide obtained by chemical synthesis is understood to mean a DNA fragment generated by automated synthesis, for example with the aid of a suitable automated apparatus.

For the present invention, the term "highly stringent conditions" is used in the sense given by Maniatis et al. 1982 (Molecular Cloning, A Laboratory Manual, Cold Spring Harbor CSH, N.Y., USA) or one of its recent editions. By way of example, the hybridization conditions are such that three washes at 65°C in the presence of 0.2 SSC, and 0.1% SDS are necessary in order to eliminate the nonhybridized fragments.

The "specific" character of the expression means that the activity of the promoter is significantly considerably higher in the cells of the cardiac tissue. Although a nonspecific expression can exist in other cells, the corresponding level of activity remains most generally very low (negligible) compared with that observed in the cardiac cells, generally lower by a factor of at least 10.

The results presented in the examples show, in this regard, a difference in expression which may reach a factor of 1000, which reflects the high selectivity of the polynucleotides according to the invention in cardiac cells *in vivo*.

Moreover, the results presented in the examples below clearly show that the use of the polynucleotides offers the system of the invention high levels of expression, above those for other promoters known to be specific for the cardiac tissue, it being possible for the difference

to exceed a factor of 100. These elements therefore illustrate the advantages and unexpected properties of the polynucleotide according to the invention in terms of strength and specificity, for the expression of nucleic acids of interest in the cardiac tissue.

Advantageously, the polynucleotide according to the invention comprises a portion of the sequence between -2266 and +92, relative to transcription position +1, whose sequence SEQ ID NO: 1 is given in the annex.

The subject of the present invention is therefore the sequences hybridizing, under high stringency conditions, with the sequence SEQ ID NO: 1.

The present invention is nevertheless not restricted to the polynucleotides containing fragments upstream of the mouse gene but relates to any functional variant or any other sequence of any other species having the same properties, namely being capable of specifically inducing expression *in vivo* of a transgene in the cardiac tissue.

Thus, persons skilled in the art will be advantageously able to refer to the sequence upstream of the human gene deposited in GenBank under the reference AF131884, whose sequence SEQ ID NO: 2 is given in the annex. The present invention thus covers any sequence comprising fragments of the sequences upstream of the gene for the CARP protein, modified for example by deletion of certain structures and which preserve identical or similar functions to that of the sequence SEQ ID NO: 1.

Preferably, the polynucleotide according to the invention exhibits at least 80%, more preferably at least 90%, identity with the sequence SEQ ID NO: 2.

The expression functional variant is understood to mean any modified sequence preserving the properties of the polynucleotides as mentioned above. The modifications may comprise one or more additions, mutations, deletions and/or substitutions of nucleotides in the sequence considered. These modifications may be introduced by conventional molecular biology methods, such as in particular site-directed mutagenesis, or in a more practical manner by artificial synthesis of the sequence in a synthesizer. The variants obtained are then tested for their capacity to control a specificity of expression in the cardiac muscle cells comparable to that of a polynucleotide having the sequence SEQ ID NO: 1.

Another subject of the invention relates to an expression cassette comprising a polynucleotide as defined above, operably linked to a transgene such that the expression of the latter is specifically directed in the cardiac muscle.

Advantageously, the cassette of the invention comprises, in addition, a signal for termination of transcription, placed in 3′ of the nucleotide sequence of the transgene.

Preferably, the transgene comprises a nucleic acid of therapeutic interest encoding a protein or an RNA which may be involved in cardiac pathologies such as cardiac insufficiency, cardiac hypertrophy, hypoxia, ischaemia or in cardiac transplant rejection.

As protein of therapeutic interest, there may be mentioned, *inter alia*:

- proteins inducing angiogenesis, such as for example members of the VEGF family, members of the FGF family and more particularly FGF1, FGF2, FGF4, FGF5, angiogenin, EGF, TGFα, TGFβ, TNFα, Scatter Factor/HGF, members of the angiopoietin family, cytokines and in particular interleukins including IL-1, IL-2, IL-8, angiotensin-2, plasminogen activator (TPA), urokinase (uPA), the molecules involved in the synthesis of active lipids (prostaglandins, Cox-1);

- proteins involved in the control of cardiac contractility, such as phospholamban, phospholamban inhibitors, SERCA-2a, β2-adrenergic receptor or dystrophin or minidystrophin (FR 91 11947);

- proteins with cryoprotective activity, which in particular block apoptosis, such as proteins which are members of the bcl family, and protein kinases such as AKT/PKB;

- transcription factors, such as for example natural or chimeric nuclear receptors, comprising a DNA-binding domain, a ligand-binding domain and a transcription activating or inhibiting domain, such as for example the fusion proteins tetR-NLS-VP16, the fusion proteins derived from oestrogen receptors, the fusion proteins derived from steroid hormone receptors, the fusion proteins derived from progesterone receptors, the proteins of the CID (Chemical Inducer of Dimerization) system described by Rivera et al., (Rivera et al., *Nature Medicine*, **2** (1996) 1028-1032). There may be mentioned in particular, as chimeric nuclear receptor, the nuclear receptors PPAR (Peroxisome Proliferator Activated Receptor) and in particular PPARγ2, as described in Applications

WO 96/23884 and FR 99 07957 and by Frohnert et al., (*J Biol Chem* **274** (1999) 3970-3977), and Mukherjee et al., (*J Biol Chem* **272** (1997) 8071-8076) either in its native form, without modification of the primary structure, or a modified PPARγ2 comprising one or more ligand-binding sites or E/F domains (Schoonjans et al. *Biochim. Biophys. Acta.* **1302** (1996) 93-109), such as PPARγ2γ2 having the sequence SEQ ID NO: 3;

- immunosuppressors such as for example interleukins 2 and 10 which make it possible to completely or partially inhibit an immune signalling pathway and thus to extend the duration of cardiac transplants;

- proteins involved as agent for reducing hypoxia such as NOS (nitric oxide synthetase), B-cell leukaemia/lymphoma 2 (bcl-2), superoxide dismutase (SOD) and catalase.

As RNA of therapeutic interest, there may be mentioned for example antisense RNAs which are useful for controlling the expression of genes or the transcription of cellular mRNAs, thus blocking translation into a protein according to the technique described in Patent EP 140 308, as well as the ribozymes which are capable of selectively destroying target RNAs as described in EP 321 201.

It is understood that the present invention is not limited to the specific examples of proteins or RNAs, but that it can be used by persons skilled in the art for the expression of any nucleic acid in cardiac cells, by simple customary experimentation operations.

The subject of the present invention is in addition vectors containing the polynucleotide or the expression cassette according to the invention. Such a vector may contain any other DNA sequence necessary for the expression of the transgene in the target tissues, and in particular may contain a replication origin which is effective in the cardiac cells.

The vector of the invention may be of a varied, in particular plasmid, episomal, chromosomal, viral or phage, nature and/or origin. It is preferably a plasmid or a recombinant virus.

By way of illustrations of the plasmids comprising a polynucleotide or an expression cassette, there may be mentioned for example the plasmids pXL3634, pXL3728 and pXL3759 which are described later.

According to a first embodiment, the vectors according to the invention are of the plasmid type. As plasmid vector, there may be mentioned, *inter alia*, any cloning and/or expression plasmids known to a person skilled in the art and which generally comprise a replication origin. There may also be mentioned new-generation plasmids carrying replication origins and/or markers which have been refined, as described for example in Application WO 96/26270.

According to a preferred embodiment, the plasmid vector is a miniplasmid and comprises a replication origin whose functionality in the host cell requires the presence of at least one protein which is specific and foreign to the said cell. Such vectors are in particular described in Application WO 97/10343.

According to a second embodiment, the vectors

according to the present invention are viral vectors. Among the latter, there may be mentioned, *inter alia*, recombinant adenoviruses, recombinant adeno-associated viruses, recombinant retroviruses, lentiviruses, herpesvirus, and vaccinia virus, whose preparation may be carried out according to methods known to persons skilled in the art. Preferably, chimeric viral vectors are used such as the adenovirus-retrovirus chimeric vectors which are described *inter alia* in Application WO 95/22617, as well as the episome/adenovirus vectors which are described by Leblois et al. (*Mol Ther* (2000) **1(4)**, 314-322) and in Application WO 97/47757.

When adenoviruses are used according to this embodiment, these are preferably vectors derived from defective adenoviruses, that is to say that they are incapable of autonomously replicating in the target cell. The construction of these defective viruses as well as their infectious properties have been widely described in the literature (see in particular S. Baeck and K.L. March, *Circul. Research*, **82**, (1998) 295-305); T. Shenk, B.N. Fields, D.M. Knipe, P.M. Howley et al. (1996), Adenoviridae: Viruses and Replication (in virology) 211-2148, EDS - Ravens publishers Philadelphia; Yeh, P. et al. *FASEB* **11** (1997) 615-623).

Various adenovirus serotypes, whose structure and properties vary somewhat have been characterized. Among these serotypes, use is preferably made in the context of the present invention of the type 2 or 5 human adenoviruses (Ad 2 or Ad 5) or adenoviruses of animal origin such as those described in Application FR 93 05954 or adenoviruses of mixed origin. Among the adenoviruses of

animal origin which can be used in the context of the present invention, there may be mentioned the adenoviruses of canine, bovine, murine (Beard et al., *Virology* **75** (1990) 81), ovine, porcine, avian or simian origin. Preferably, the adenovirus of animal origin is a canine adenovirus, more preferably a CAV2 adenovirus (Manhattan or A26/61 strain) as described in Application WO 94/26914.

The defective adenoviruses of the invention comprise in general an inverted terminal repeat (ITR) at each end, a sequence allowing encapsidation (Psi), the E1 gene and at least one of the genes E2, E4 and L1-L5 having been moreover inactivated by any technique known to persons skilled in the art (Levero et al., *Gene*, **101** (1991) 195, EP 185 573; Graham, *EMBO J.* **3** (1984) 2917).

Advantageously, the recombinant adenovirus used in the context of the invention comprises a deletion in the E1 region of its genome. More particularly still, it comprises a deletion of the E1a and E1b regions. By way of precise example, there may be mentioned deletions affecting nucleotides 454-3328, 382-3446 or 357-4020 (with reference to the genome of Ad5).

According to a preferred variant, the recombinant adenovirus used in the context of the invention comprises, in addition, a deletion in the E4 region of its genome. More particularly, the deletion in the E4 region affects all the open reading frames. There may be mentioned, by way of precise example, the 33466-35535 or 33093-35535 deletions. Other types of deletions in the E4 region are described in applications WO 95/02697 and WO 96/22378, which are incorporated into the present by way of reference.

As regards the adeno-associated viruses (AAV), they are relatively small-sized DNA viruses which integrate into the genome of the cells which they infect, in a stable and site-specific manner. They are capable of infecting a broad spectrum of cells, without inducing any effect on cell growth, morphology or differentiation. Moreover, they do not appear to be involved in pathologies in humans. The AAV genome has been cloned, sequenced and characterized. It comprises about 4700 bases and contains, at each end, an inverted terminal repeat (ITR) of about 145 bases, serving as replication origin for the virus. The remainder of the genome is divided into 2 essential regions carrying the encapsidation functions: the left portion of the genome, which contains the rep gene involved in viral replication and the expression of the viral genes; the left portion of the genome, which contains the cap gene encoding the virus capsid proteins.

The use of AAV-derived vectors for the transfer of genes *in vitro* and *in vivo* has been described in the literature (see in particular WO 91/18088; WO 93/09239; US 4,797,368, US 5,139,941, EP 488528). These applications describe various AAV-derived constructs in which the rep and/or cap genes have been deleted and replaced with a gene of interest, and their use for transferring *in vitro* (on cells in culture) or *in vivo* (directly in an organism) the said gene of interest. The defective recombinant AAVs according to the invention may be prepared by co-transfection, into a cell line infected with a human helper virus (for example an adenovirus), of a plasmid containing the nucleic sequences of the invention bordered by two AAV inverted terminal repeats (ITR), and of a

plasmid carrying the AAV encapsidation genes (rep and cap genes). The recombinant AAVs produced are then purified by conventional techniques.

Lentiviruses may also be used according to this embodiment; they allow the transfer and efficient and stable integration of a gene of interest into quiescent cells.

There may be mentioned for example HTLV-1 of animal lentiviruses such as FIV (feline infections virus), EIAV (equine infectious anemia virus; WO 98/51810), BIV (bovine immunodeficiency virus), SIV (simian immunodeficiency virus), CAEV (caprine arthritisencephalitis virus) (WO 98/39463; Naldini et al. *Science* **272** (1996) 263-267; Schnele et al. *Hum Gen Ther* **11** (2000) 439-447), or a lentivirus related to the one which causes AIDS, HIV-2 which is not highly pathogenic in humans (Kundra et al., *Hum Gen Ther* **9** (1998) 1371-1380).

The expression cassettes may be inserted at various sites of the recombinant genome. It may be inserted at the level of the E1, E3, or E4 region, as a replacement for suppressed or surplus sequences. It may also be inserted at any other site, outside of the sequences necessary in cis for the production of the viruses (ITR sequences and encapsidation sequence).

It will be noted, however, that the introduction of the sequences according to the present invention into the vectors described above is not essential, such that the cardiac cells may be directly transfected with DNA comprising these sequences.

The nucleic sequences according to the present invention may be introduced after covalent coupling of the nucleic acid with compounds promoting their penetration into the cells or their transport to the nucleus, the resulting conjugates being optionally encapsidated into polymeric microparticles, as in International Application WO 94/27238.

According to another embodiment, the nucleic sequences may be included in a transfection system comprising polypeptides promoting their penetration into cells, as in International Application WO 95/10534.

These polynucleotides, cassettes and vectors may be administered in situ by any means known to persons skilled in the art, for example by coronary infusion (Barr et al., *Gene Ther*, **1**, (1994) 51-58), by intracardiac injection, by epicardiac injection, that is to say through the ventricular wall (Guzman et al., *Cir Res*, **73** (1993) 1202-1207), by intrapericardiac injection (Fromes et al., *Gene Ther*, **6** (1999) 683-688), or by retrofusion of the coronary veins (Boeckstegers et al., *Circulation*, **100** (Suppl I) (1999), I-815).

The polynucleotides, cassettes or vectors according to the invention may advantageously be administered in the form of a composition containing them, for example with the aid of a chemical or biochemical transfer agent facilitating their transfection into the cardiac cells. The expression "chemical or biochemical transfer agent" is understood to mean any compound facilitating the penetration of a nucleic acid into a cell. This may include cationic agents such as cationic lipids, peptides, polymers (Polyethylenimine, Polylysine), nanoparticles, or

non-cationic agents such as non-cationic liposomes, non-cationic nanoparticles or polymers, Such agents are well known to persons skilled in the art and are in particular described in applications WO 95/18863, WO 97/18185 and WO 98/15639.

The present invention, in addition, relates to medicaments containing such polynucleotides, expression cassettes or vectors as well as pharmaceutical compositions containing them in a pharmaceutically effective quantity as well as pharmaceutically compatible excipients.

Such polynucleotides, expression cassettes or vectors may be advantageously used for the manufacture of medicaments for delivering to the cardiac tissue which may express in particular a gene encoding a protein of interest, for the treatment of cardiac diseases and in particular for the treatment and/or prevention of cardiac insufficiency, hypoxia, cardiac hypertrophy, mycarditis, cardiac ischaemia, or for preventing rejection during cardiac transplant.

Such a medicament may, for example, comprise a cassette or vector according to the invention which is capable of expressing the functional form of the impaired gene according to the cardiac pathology which it is desired to treat.

Preferably, the pharmaceutical composition contains pharmaceutically acceptable vehicles for an injectable formulation, in particular for intracardiac injection. This may include in particular isotonic, sterile saline solutions (monosodium or disodium phosphate, sodium, potassium, calcium or magnesium chloride, and the like, or

mixtures of such salts), or dry, in particular freeze-dried, compositions, which, upon addition, depending on the case, of sterilized water or of physiological saline, allow the preparation of injectable solutions. Other excipients may be used, such as, for example a hydrogel. This hydrogel may be prepared any biocompatible and non-cytotoxic (homo or hetero) polymer. Such polymers have for example been described in application WO 93/08845. Some of them, such as in particular those obtained from ethylene and/or propylene oxide are commercially available. The doses used for the injection may be adjusted according to various parameters, and in particular according to the aim pursued (labelling, pathology, screening and the like), the transgene to be expressed, or the duration of expression desired.

In general, the recombinant adenoviruses according to the invention are formulated and administered in the form of doses of between 10^4 and 10^{14} pfu, and preferably 10^6 and 10^{10} pfu. The term pfu (plaque forming unit) corresponds to the infectious power of a viral solution, and is determined by infecting an appropriate cell culture, and measuring the number of plaques of infected cells. The techniques for determining the pfu titre of a viral solution are well known in the art.

The subject of the present invention is, in addition, a method of expressing a transgene of therapeutic interest during which the polynucleotides, cassettes or vectors according to the present invention are used, such that the transgene can be expressed.

Moreover, the invention also relates to any cell modified with a cassette or a vector (in particular an adenovirus) as described above. The expression "modified" cell is understood to mean any cell containing the polynucleotide or the cassette according to the invention. These cells may be intended for implantation into an organism, according to the methodology described in application WO 95/14785. These cells are essentially human cardiac cells.

The present invention also relates to transgenic animals and in particular mice carrying a polynucleotide or a cassette as defined above in which the gene encoding the protein of therapeutic interest is replaced with a reporter gene. Such transgenic mice may be used for screening molecules for their activity on the regulatory sequences of the gene encoding the CARP protein.

Molecules may be administered to mice, and then after sacrificing, histological sections are prepared in order to identify the tissues stained with the reporter gene.

The transgenic animals according to the present invention also constitute molecular biology study means for understanding molecular mechanisms underlying cardiac pathologies of genetic origin, such as cardiac insufficiency, cardiac hypertrophy, cardiac hyperplasia and myocardial infarction.

By way of example, there may be mentioned murine models for studying myocarditis in which the gene encoding interferon-1 (IFN-1) is inactivated (Aitken et al., *Circulation*, **90** (1994) 1-139).

Other animal models of interest according to the present invention may comprise the polynucleotide according to the invention linked to transgenes such as protooncogenes or oncogenes, for example c-myc, thus constituting models of hyperplasia (Jackson et al., *Mol Cell Biol*, **10** (1990) 3709-3716), p21-ras for models or ventricular hypertrophy (Hunter et al., *J Biol Chem*, **270** (1995) 23176-23178), nuclear antigen of the Epstein-Barr virus for studying certain cardiomyopathies (Huen et al., *J Gen Virol*, **74** (1993) 1381-1391).

According to another embodiment, the transgenic animals according to the invention constitute experimental models of cardiac hypertrophy and comprise an expression cassette in which the transgene encodes for example calmodulin (Gruver et al., Endocrinology, 133 (1993) 376-388), interleukin-6 or the interleukin-6 receptor (Hirota et al., *Proc Natl Acad Sci*, **92** (1995) 4862-4866), cardiotrophin-1 (Pennica et al., *Proc Natl Acad Sci*, **92** (1995) 1142-1146), and finally the α-adrenergic receptor (Milano et al., *Proc Natl Acad Sci*, **92** (1994) 10109-10113).

Additionally, the polynucleotides according to the invention, modified such that they allow an increase in the expression of the CARP gene, also form part of the invention. The transgenic animals thus obtained thus constitute experimental tools for myocardial infarction (Stanton et al., *Circul Res*, **86** (2000) 939-945).

To carry out the present invention, a person skilled in the art can advantageously refer to the following manual "SAMBROOK et al. (Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York

1989), or one of its recent editions.

The present invention is described in greater detail with the aid of the following examples which should be considered as illustrative and nonlimiting.

LEGEND TO THE FIGURES

Figure 1: illustrates the nucleotide sequence (SEQ ID NO: 1) of the polynucleotide upstream of the gene encoding the mouse CARP protein;

Figure 2: illustrates the nucleotide sequence (SEQ ID NO: 2) of the polynucleotide upstream of the gene encoding the human CARP protein;

Figure 3: is a schematic representation of the plasmid pXL3634;

Figure 4: is a schematic representation of the plasmid pXL3728;

Figure 5: illustrates the relative activity *in vitro* of the plasmids pXL3635 and pXL3634, with respect to the reference activity of the CMV promoter (pRL-CMV). The activity of each promoter represents the Photinus pyralis luciferase activity normalized with the Renilla reniformis luciferase activity.

Figure 6A: is a schematic representation of the plasmid pXL3759;

Figure 6B: is a schematic representation of the adenovirus AV1.0 CARP-Luc+;

Figure 7A: illustrates the luciferase activity (pg luciferase/heart) 7 days after intracardiac transdiaphragmatic injection in rats of variable quantities of plasmids pXL3031 and

pXL3634;

Figure 7B: illustrates the luciferase expression (pg luciferase/heart) 7 days after intracardiac transdiaphragmatic injection in rats hearts of 25 µg of plasmids pXL3031 and pXL3635, pXL3130, and pXL3153.

Figure 8: represents the ratio of the expression of Luciferase in the heart relative to the expression in the muscle as a function of the expression in the heart obtained following intracardiac administrations of plasmids pXL3031, pXL 3634, pXL3635, pXL3153, and pXL3130.

EXAMPLES

Example 1: Characterization of the polynucleotide upstream of the CARP gene

A BamHI-XhoI fragment of 2.3 Kb of the sequence in 5' of the mouse gene encoding the CARP protein was cloned and sequenced on both strands according to the chain termination method (Sanger et al., 1977, PNAS, 74, 5463) using the Sequenase kit (United States Biochemical, Cleveland, Ohio). The sequence is represented in Figure 1, and therefore comprises a portion upstream of the gene encoding the mouse CARP protein between nucleotides -2266 and +92 relative to transcription position +1 (SEQ ID NO:1).

Example 2: Construction of CARP plasmid vectors

2.1 Plasmid pXL3634

The BamHI-XhoI fragment of 2.3 Kb characterized in Example 1 was cloned after filling of the BamHI site in the plasmid pGL3-Basic (Promega) previously digested with XhoI and SmaI, in order to obtain the plasmid pXL3634. A schematic representation of this plasmid is presented in Figure 3.

2.2 Plasmid pXL3728

The plasmid pXL3728 is obtained from the plasmid pXL3179 which is a vector derived from the plasmid pXL2774 (WO 97/10343) in which the gene encoding a fusion between the signal peptide of human fibroblast interferon and the cDNA of FGF1 (Fibroblast Growth Factor 1) (sp-FGF1, Jouanneau et al., *PNAS* **88** (1991), 2893-2897) was introduced under the control of the promoter obtained from the human cytomegalovirus early region (hCMV IE) and the polyadenylation signal of the SV40 virus late region (GenBank SV4CG).

The BamHI-XhoI fragment of 2.3 Kb characterized in Example 1, whose ends have been filled, was cloned into the plasmid pXL3179 (pCOR CMV-FGF), previously digested with XbaI and EcoRI, in order to obtain the plasmid pXL3728. A schematic representation of this plasmid is presented in Figure 4.

2.3 Plasmid pXL3729

An EcoRI-SalI fragment of the plasmid pXL3634 was cloned into the plasmid pXL3728 previously digested with EcoRI-SalI in order to obtain the plasmid pXL3729.

Example 3: Comparative plasmids

3.1 Plasmids pXL3130 and pXL3153

Plasmids pXL3130 and pXL3153 contain respectively the human smooth muscle α- actin promoter (-680,+30) and the mouse SM22 promoter (-436,+43) coupled to CMV enhancer (-522, -63) as described in application WO 00/18908.

3.2 Plasmid pXL3635

The RSV -229,+34 promoter was cloned from a construct containing a longer version of the RSV promoter (contained in Ad1.0RSVLAcZ, Stratford-Perricaudet et al., *J Clin Invest* **90** (1992) 626-30) by PCR by means of the primers 5'-GGC GAT TTA AAT AAT GTA GTC TTA TGC AAT-3' and 5'-GGG GTC TAG AAG GTG CAC ACC AAT GTG GTG A-3' which introduce, respectively, an SwaI and XbaI site in 5' and 3' of the PCR fragment. These two restriction sites were then used to introduce the promoter fragment into pGL3-basic to generate pXL3635.

3.2 Plasmid pXL3031

The plasmid pXL3031 is as described by Soubrier et al., *Gene Ther.* **6** (1999), 1482-8. It is a vector derived from the plasmid pXL2774 (WO 97/10343) in which the luc gene encoding the modified Photinus pyralis luciferase (cytoplasmic) obtained from pGL3basic (GenBank: CVU47295) was introduced under the control of the promoter obtained from the human cytomegalovirus early region (hCMV IE, GenBank HS5IEE) and of the polyadenylation signal of the SV40 virus late region (GenBank SV4CG).

Example 4: Cell cultures

Primary cultures of rat cardiomyocytes were established. For that, gestating rats were killed in a chamber saturated with CO_2. After opening the abdomen, the uterine horns are removed and washed in PBS at room temperature. The embryos are released from their envelopes and the placenta cut (10 to 12 embryos per rat). The hearts are removed and washed in ADS/glucose. Under a binocular lens, the auricles and large vessels are removed, and then the hearts are again cleaned in ADS/glucose so as to retain only the ventricles and rinsed 3 times in sterile ADS/glucose.

The hearts are then trypsinized in 0.3 ml of an ADS/glucose/trypsin mixture per heart, using trypsin T 4674 (Sigma, St Louis, Missouri) at a final concentration of 0.1 mg/ml, for 20 min at 37°C, with gentle stirring (60 to 100 revolutions per min).

The supernatant is then removed, and trypsin inactivated by addition of 1 ml of decomplementized FCS. After centrifugation at 1500 rpm for 10 minutes, the supernatant is removed and the cardiac cells are taken up in 1 ml of decomplementized FCS. In parallel, the steps of treating with trypsin are repeated 5 to 6 times until complete dissociation of the cells is obtained. The pool of cells is centrifuged at 1500 rpm for 10 minutes, then washed twice in FCS and the cells are finally filtered on a grid filter.

The cells thus separated are then placed in culture at a concentration of 10^6 cells/well for a 24-well plate or at a concentration of 2×10^6 cells/well for a 12-well plate. Each well contains 1 ml of culture medium.

The culture medium comprises, for a total volume of 100 ml, 68 ml of DMEM (without pyruvate) (Gibco-BRL), 17 ml of M199 (Sigma M 4530), 10 ml of decomplementized horse serum (Sigma H6762), 5 ml of decomplementized FCS (Gibo-BRL) and 1 ml of 100X Peni/Strepto/glutamine mixture (Gibco-BRL).

The cardiomyocytes are cultured for a period of about 1 or 2 days.

Example 5: Transfection of primary cultures of cardiomyocytes

The primary cultures of cardiomyocytes are cotransfected with a total quantity of DNA equal to 500 ng per well, comprising 1 ng of a plasmid pRL-CMV (Promega Inc., Madison, WI), variable quantities ranging from 1 to 100 ng of each of the plasmids pXL3635 and pXL3634 as described above, qs 500 ng of pUC19.

For that, the mixture of the plasmids is incubated with 6 nmol of RPR 120535B (Byk et al., *J Med Chem.* **41** (1998) 229-35) per µg of DNA (0.3 µl of solution of lipid at 10 mM) in a final volume of 20 µl in 150 mM NaCl, 50 mM bicarbonate, and then vortex-mixed for 5 seconds, and again incubated for about 20 to 30 minutes at room temperature.

The mixture is then added to 250 µl of serum-free medium and incubated with the cells for at least 2 hours. The medium is finally removed and the cells are incubated for a period ranging from 24 hours to 7 days at a temperature of 37°C in the presence of 5% CO_2.

The cells were harvested at 24 hours or at 48 hours after transfection and the Renillia luciferase and Firefly

luciferase activities are analysed with the Promega Dual Luc kit according the manufacturer's instructions. The activities are read on a Victor apparatus.

Example 6: Comparative evaluation of the *in vitro* activity of the polynucleotide

The relative activities of the CARP polynucleotide (pXL3634) and of the RSV (pXL3635) promoters were evaluated *in vitro* in transit transfection in primary cultures of rat cardiomyocytes and were expressed relative to the activity of the plasmid pRL-CMV (Figure 5).

The results show that the polynucleotide upstream of the CARP gene used (pXL3634) has a very low *in vitro* activity, of the order of 0.04% relative to that of the CMV promoter.

The relative activity of the nonspecific strong RSV promoter (pXL3635) are also low, respectively of the order of 0.05% and 0.68% of that of the reference CMV promoter.

Example 7: Construction of an adenovirus

An adenovirus allowing the expression of the luciferase under the control of the CARP promoter was constructed according to the method of Crouzet et al. (*PNAS*, **94** (1997) 1414-1419), the expression cassette being identical to that of the plasmid pXL3634 (Figure 3).

A shuttle vector allowing recombination in *E.coli* was constructed in two stages. First of all, the CARP promoter (fragment: XhoI filled with Klenow/BamHI) was introduced in pXL3474 (digested with ScaI and BglII) between the regions ITR-Ψ and pIX in order to generate the plasmid pXL3758. pXL3759 was then generated by introducing into

pXL3758, digested with BstBII (filled with Klenow) and BstEII, the fragment containing the luciferase cDNA and the SV40 polyadenylation site (BamHI fragment filled with Klenow/BstEII of pXL3634). pXL3759 is schematically represented in Figure 6A.

Homologous double recombination in E. coli was accomplished as described above, against a plasmid pXL3215 containing an ΔE1/ΔE3 adenoviral genome in which an RSV-LacZ expression cassette is introduced into the E1 region. The plasmid pXL3215 is a derivative of the plasmid pXL2689 which contains the replication origin of the plasmid RK2, the tetracycline resistance gene (Crouzet et al. PNAS, 1997). The product of this double recombination, the plasmid pXL3778, is checked by sequencing of the expression cassette. After cleavage with PacI in order to release a linear viral genome, the plasmid is transfected into the Per.C6 cell line (WO 97/00326) in order to generate the virus AV1.0CARP-Luc+.

The virus is also checked by sequencing of the expression cassette by restriction analysis and the presence of RCA E1+ (replication competent adenovirus) particles is tested for by hybridization with a probe Ψ.

Stocks with high virus titre are obtained by amplification of the virus in the Per.C6 line and the viral particles purified on a CsCl gradient. The titre of this virus in viral particles/ml (vp/ml) is obtained by chromatography and its activity is checked *in vitro* by titration of the luciferase activity after infection of skeletal or cardiac muscle cells and comparison with viruses used as a control comprising a CMV promoter.

Example 8: Injection of DNA *in vivo*

CD SPRAGUE rats weighing 200 g are anaesthetized with a Ketamine (70 mg/ml)/Xylazine (6 mg/ml) mixture at 1 ml/kg injected by the intraperitoneal route.

The intramyocardiac injections are carried out, after laparotomy, by the transdiaphragmatic route with a 100 µl Hamilton glass syringe connected to a Steriflex catheter (ref. 167.10 G19 V) provided with a stop flange and ending with a BD 26G*3.8 needle (short bezel).

50 µl of the DNA solution, adjusted to 0.9% of NaCl, are thus injected over 5 seconds.

After sacrificing the animals, the hearts are removed, rinsed in a 0.9% NaCl solution and macroscopically examined. They are then analysed for the luciferase activity by means of a kit (Promega E151A) after grinding with the aid of a homogenizer (Ultra-thurax, Diax600 Heidolph) in a lysis buffer in the kit supplemented with protease inhibitors (CØmplete™, Roche Diagnostics), followed by centrifugation for 20 minutes at 4000 rpm at 4°C. The readings are made on the apparatus: LUMAT LB 9501 (10 µl of supernatant + 50 µl of Promega luciferase substrate). Luciferase activities are converted to luciferase mass per heart (pg luciferase/heart) using calibration described in Mir et al (*PNAS* **96** (1999), 4262-4267).

Alternatively, the hearts are fixed in 3.7% paraformaldehyde and analysed by immunohistochemistry for the expression of FGF-1.

Example 9: Comparative evaluation of the *in vivo* activity of the CARP polynucleotide

The results assembled in Figure 7A show that the levels of expression of luciferase obtained upon injection of increasing doses 1, 5, 25 and 125 µg of plasmids pXL3031 and pXL3634, are not significantly different, thus clearly demonstrating that the polynucleotide upstream of the CARP gene is capable of inducing high levels of expression equivalent to those of a strong promoter such as CMV.

On the other hand, the expression obtained with another strong viral promoter, the RSV promoter (pXL3635) is weaker than that obtained with the CMV promoter or the polynucleotide upstream of the CARP gene (Figure 7B).

Moreover, the addition of the CMV enhancer upstream of smooth muscle cell promoters (SM α-actin, pXL3130 or SM22, pXL3153) although demonstrated to be highly efficient in vitro (WO 00/18908) appears to be ineffective in cardiac cells *in vivo*.

Example 10: Evaluation of the specificity of expression of the CARP polynucleotide

25 µg of each of the plasmids pXL3634, pXL3435 and pXL3031 were administered to rats by intracardiac transdiaphragmatic injection.

In parallel, intramuscular injections were performed into the cranial tibial muscle of groups of mice with 10 µg of each of these plasmids with or without electrotransfer.

The expression of luciferase was analysed 7 days after the injection as described (*PNAS* **96** (1999), 4262-

4267).

The levels of expression of luciferase in the heart were expressed relative to the levels observed in the cranial tibial muscle, and are assembled in Figure 8.

The results clearly show that the polynucleotide upstream of the CARP gene and the CMV promoter are the only two promoters capable of inducing the highest expression in the cardiac tissue. However, the heart/muscle expression ratio is 1 with the CMV promoter, whereas this ratio is close to 100 when the polynucleotide upstream of the CARP gene is used, which clearly shows the very high selectivity of the latter for the cardiac tissue.

The superiority of the specificity of the expression of the polynucleotide is also clear relative to other constructs comprising an enhancer and a promoter specific for smooth muscle cells such as that of the gene coding for the protein SM-22 and for actin for which the heart/muscle expression ratios are also presented in Figure 8 by way of illustration.

CLAIMS

1. Polynucleotide, characterized in that it comprises a fragment of the sequence upstream of the coding part of the gene for the CARP protein having the sequence SEQ ID NO: 1, or of a sequence hybridizing under high stringency conditions with the said sequence, the said polynucleotide being capable of inducing a specific expression *in vivo* of a gene operably linked to the said polynucleotide, in cardiac cells.

2. Polynucleotide, characterized in that it exhibits at least 93% identity with the sequence SEQ ID NO: 1.

3. Polynucleotide according to Claim 1, characterized in that the said fragment is contained in the sequence situated between nucleotides -2266 and +92 of the mouse gene for the CARP protein (SEQ ID NO: 1).

4. Expression cassette, characterized in that it comprises a sequence encoding a protein or an RNA, of therapeutic interest, placed under the control of a DNA sequence as defined in any one of Claims 1 to 3.

5. Expression cassette, characterized in that it comprises a sequence encoding a protein or an RNA, of therapeutic interest, placed under the control of a DNA sequence exhibiting at least 80% sequence identity with the sequence SEQ ID NO: 2.

6. Cassette according to either of Claims 4 and 5, characterized in that the said protein, or the said RNA is capable of activating the growth of the cardiac cells, of reducing or suppressing an immune response, of inducing angiogenesis, of correcting muscle contractility, cardiac hypertrophy, cardiac insufficiency and myocarditis.

7. Cassette according to one of Claims 4 to 6, characterized in that the said protein of therapeutic interest is a protein of the family comprising VEGF, FGF, angiopoietins, and cytokines.

8. Cassette according to one of Claims 4 to 6, characterized in that the said protein of therapeutic interest is an activating or inhibiting transcription factor.

9. Cassette according to one of Claims 4 to 6, characterized in that the said protein of therapeutic interest is an immunosuppressive protein such as interleukin-10, interleukin-2 and interleukin-8.

10. Cassette according to one of Claims 4 to 6, characterized in that the RNA of therapeutic interest is an antisense RNA which makes it possible to control the expression of genes or to block the transcription of mRNA in cardiac cells.

11. Cassette according to Claim 4 to 6, characterized in that the said protein is an agent for reducing hypoxia, chosen from NOS (nitric oxide synthetase), superoxide dismutase (SOD) and catalase.

12. Vector, characterized in that it contains a polynucleotide according to one of Claims 1 to 3.

13. Vector, characterized in that it contains an expression cassette according to one of Claims 4 to 11.

14. Vector, characterized in that it contains a replication origin which is effective in cardiac cells.

15. Vector according to one of Claims 12 to 14, characterized in that it is a plasmid, a cosmid or any DNA not encapsidated by a virus.

16. Vector according to either of Claims 12 and 14, characterized in that it is a recombinant virus chosen from an adenovirus, a retrovirus, a herpesvirus, an adeno-associated virus or a derivative thereof.

17. Composition, characterized in that it contains a vector according to one of Claims 12 to 16.

18. Composition comprising a vector according to one of Claims 12 to 16 and a chemical or biochemical transfer agent.

19. Medicament, characterized in that it contains a vector according to one of Claims 12 to 16.

20. Pharmaceutical composition, characterized in that it contains an effective quantity of a polynucleotide or of a vector according to one of Claims 1 to 3 and 12 to 16.

21. Use of a polynucleotide according to one of Claims 1 to 3 or of a vector according to one of Claims 12 to 16 for the manufacture of a medicament intended for the treatment of cardiac insufficiency.

22. Use of a polynucleotide according to one of Claims 1 to 3 or of a vector according to one of Claims 12 to 16 for the manufacture of a medicament intended for the treatment of cardiac hypertrophy.

23. Use of a polynucleotide according to one of Claims 1 to 3 or of a vector according to one of Claims 12 to 16 for the manufacture of a medicament intended for the treatment of hypoxia.

24. Use of a polynucleotide according to one of Claims 1 to 3 or of a vector according to one of Claims 12 to 16 for the manufacture of a medicament for preventing rejection during a cardiac transplant.

25. Transgenic animal, characterized in that it carries a polynucleotide according to one of Claims 1 to 3, in which the gene encoding the protein of therapeutic interest is replaced with a reporter gene.

26. Method of expressing a gene of therapeutic interest *in vivo*, characterized in that

a) a vector according to any one of Claims 12 to 16 is isolated, and

b) an effective quantity of the said vector is introduced in the cardiac tissue, under conditions such that the said gene of interest is expressed.

FIGURE 1

```
   1 ggatcctttc atgtttaaca atatcaaccc taacccaagg ggaacagcct gcctgacagt
  61 ggctttgcca cccatgaata cttcctagtc tagtccgttt gtgaaactca gcccatccca
 121 acacttctgc aagccccatc ctctacaagg tgctcattgg gaatttcctg gagcttctct
 181 ttcaggatca gcctgattct agggcagcag ttctcaacct gggggcctcg acccctttgg
 241 gggaatcaaa cgacccttta caggggtcac atatcatcta tcctatatgt caggtattta
 301 cattacgatt cgtaacagta gcaaaattac aggtatgaaa tagcaatgaa ataattttat
 361 gattgaaggt caccacaaca tgaggccgcc acactgttct agagaaaaat cacctgggtg
 421 gggaaaggtt tgggaaagcc tttctgtcca ttcttcattc ttcaaagtga tgtgttcaca
 481 gaaagccttt cagctgttct gctggggctc ttagtaagtc tgagtaggaa ctgtatgtac
 541 caggtctgct tcttatgggt ggagccaaga cgcatcgtgg gtggagcgaa gacgcaacct
 601 caccttctag ctctgcatcc atagcaagta gcctaatgtt tctgtgtcta ggtgtcatct
 661 ctgtgaatcg agatccttgg ccttgcttga attagggagg cacaaaatac tcagagattc
 721 aagactgctc agcagcccag agtccttcct caaaggaaag gtctcaactc tcagcccccc
 781 ttagctctga gtcaggcctg gaacaaacgg ccacaggaat gagaaaagct gccatagctg
 841 cttgtcactt caagaggtca aagaaaatag tgttaaccat gaaaacgaga agaccaacag
 901 ttatccattg atagcgtctc aggacagata ggacagagag aacactagga gaggggaacc
 961 cacgaaggac aaggtattag tgtgttggtt ttcagggcaa tgtcttgtac tgaagattct
1021 agaaacacaa tttgctggtt gaacagctga agtggggtgg gggttcttac cccatgttca
1081 tggaagggtg agtgaggaga gacagatata tgatggccag cataacaaac atacacaaca
1141 ccctaattaa cacttccctc ttctactgac acccccttca ctctcctctt tcataaaaaa
1201 taaaaaaagt attttatgtg gctcttacga tagaatcttt cctcgaacta taaaaagatc
1261 taaatattta tatttttcac attttaatat cttagcgatg acaagccaga aacaagtatt
1321 ttttgcctct ctcaacagca aagcttgggg ccttttttgtt tccgtgttag gaatagaaca
1381 cgagagcccc gtgtatctag gcagatgctc tatcattagc ccatgagtct ccagcctcag
1441 acgcacattt ttctcgggct ctcttaagct tttcccacag cattgggaaa ctttactgac
1501 agcatccaag ttgtgcttct gctaagaact ggactcacat ctctctgtgc atcacttcgg
1561 cccgttttgg ggtagatcct ctgattagcc ttcagattta gaacacggtg agcctgtggt
1621 gcactaatta tggccagtga caccatagag tcaaagtgca ttactgaatg ctttcaattt
1681 ctcctaatgc tggtacgatg gcatgtcaca gggccatttt agctgcagac atcactccag
1741 agaattccaa acagatagag acaagtggca cccagaccca tctccttccc ctcgggctga
1801 ttatccccag aaataggatg tcccaaagca acacttccca gccaactgga gtgctgataa
1861 gtccagttat cagaaagata tggctgtaag tgtgatgcac agtgcttgca ttttcttgat
1921 acgttagtca tatgagagct gacaaagaag gaaaaagagc agcgatgtgg tgcaatatta
1981 acaggcagct gtcccctggc ttcccgatac gtgggatgac tcgcattgct gagcggtgtg
2041 gtcactgcca aaggaatgac cctctcacat ttcttcctga ttcgcatacg ccgcggccag
2101 cttgtcatct ccctcttggg cttcccagac actaagtctg gaatgaaaat tcacctgcct
2161 ctgaattggc cactggtggg ggcaggggtg tgacttggct tcccaggctg gaagattatc
2221 tcacccagcc ctagctatat aacgggctgg tgtggagggg ctccacaggg ccagttccag
2281 gggttcatcc acaagagaga aaaacataga ctcgaggtct agggagcttg catgcctgca
2341 ggtcggaggc caccatgg
```

FIGURE 2

```
   1 ctgcagcaag ttacttaatg ttttttgcct cagcatcctc tctgtaaaat gagagcatta
  61 gtcttgctcc aacttcgagg gcatggacag ctctgggatt tcatatccaa gacccttaaa
 121 catcccacag tccttccccc aaacacttct cctcctaata cctccctcag tttgggtcag
 181 gcctggaaca aaaaggcata cgaaatggta gaaaaagtgt ccatgactac ttctgactta
 241 gatgaagaga ccaatgaaaa tagtaatgac tctgtttgct tcagcaggac atatactaaa
 301 ataggagcta tacaaagaag attagcatgg actctgtgca agaatgacac acaaatttgt
 361 gaaacattcc atatattaaa aataaataaa taataaagag aaaaggaaaa aattaaaaag
 421 aaaatagtga tagctgtgtc catctcaaag aaaagcccag gagatttcct ttatttaccc
 481 cctttaagat agaatattag gagaccggaa catatgatac aggaggtact gggagggtcc
 541 ctctttgtca atgttttgtc ttggggtggg gagtcgatgt cttctcaaag tttcagaaac
 601 accatccact gactgagcat tcaaggggca agaggagaat ggcagccaca tttgttgatt
 661 gggtgagttt ggggagaaat agacacacaa aggtcaaaca taacttccta attaacactt
 721 ccctccattc acaattccct tctcccattc ttctctcctg tcttttacts akaraaaccc
 781 agtttttcct gaaactataa aaataccccc agtatgttta cataatttac acctcaaaga
 841 ttagaaacca gaaatagaga ccttttcaac ccttccggaa gcaaagtgca ttatccctcc
 901 agccacgtgt ctcaaatctt gatgcatcag aatcatctgg gtgctttkaa attcaagatg
 961 attcctacga gttaccataa atcaactcag aattccctgg agtggggcca gggatctgta
1021 tttctgacaa gctcccacag gtgattcctt tccccacagc atttgagaac ttcagctcaa
1081 tgacctaatc agagtcctgc cattgctaat atctggtctc atttttbtca tatatatata
1141 tagtatttgt ggtagagatg ggattttgcc atgttgccca ggctagtatt gaactcctaa
1201 gctaagcaat cttcctgtct ctgcctccca aaatgttggg attacaggtg taagccactg
1261 cacccggctg atagctggtt tcatttactc tatttcttga ccactctgat ccattttgaa
1321 gtaaaaatgc tccaattatt atgctgtttt agaacacggt aagcatgtca tgtgctaatg
1381 gccagtgaca tcataaaaga aaagtgcatt actgaatgct ttcaatgtct tataatgatg
1441 gtaaggtggc atgtcatggg gcctatttag cccagacatc actccaaaga attccaaaca
1501 gatatagaca agtgccttta gggcccagat cccttccccc caggctgttt acccagggaa
1561 taggatgtcc tgggacaagt ttcccctaag tgaagtgttg ataagtctgc ttatcagaaa
1621 gatattactg ggggtgtgat atgtagggca tctacatttt cttgataggt agtcatatga
1681 aagctgacaa agaaaaaaag ggcagtgatg tggtgcaatg tcaacagaca gctgtccccc
1741 gactcttgac aaataggatg acttgcattg ctgagcgatg tgatcaccac caaaggaatg
1801 gccctctcac atttcttcct gattcacata ttcagcaggg ttagcttgtc ctcccctccc
1861 tcttcagctt cccagacact gagtctggaa tgaaaattca cctgcctctg agttggctcc
1921 taatgggggc gggagtgtta cttcggttcc caggttggaa gattatctca cccggcccca
1981 gctatataag ctgaccggtg tggaggggcc cagcagggcc aactccaggg attccttcca
2041 cgacagaaaa acatacaaga ctccttcagc caac
```

FIGURE 3

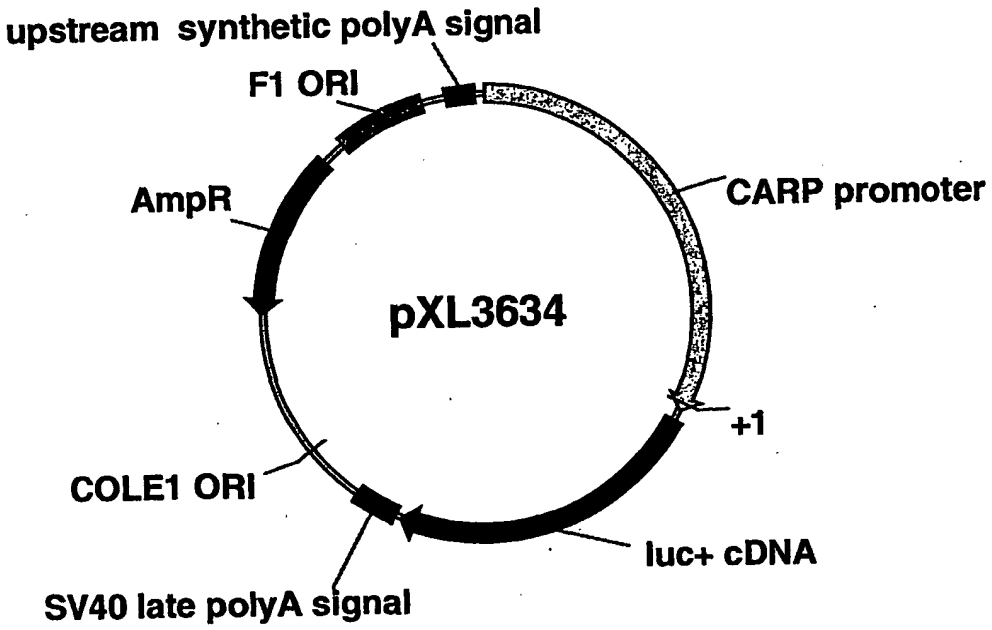

upstream synthetic polyA signal
F1 ORI
AmpR
CARP promoter
pXL3634
+1
COLE1 ORI
luc+ cDNA
SV40 late polyA signal

FIGURE 4

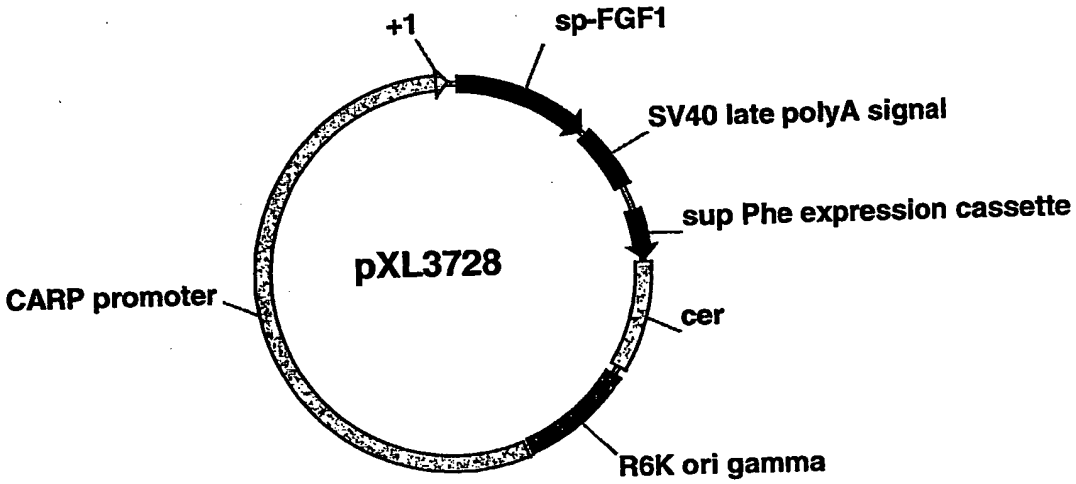

+1
sp-FGF1
SV40 late polyA signal
sup Phe expression cassette
pXL3728
CARP promoter
cer
R6K ori gamma

FIGURE 5

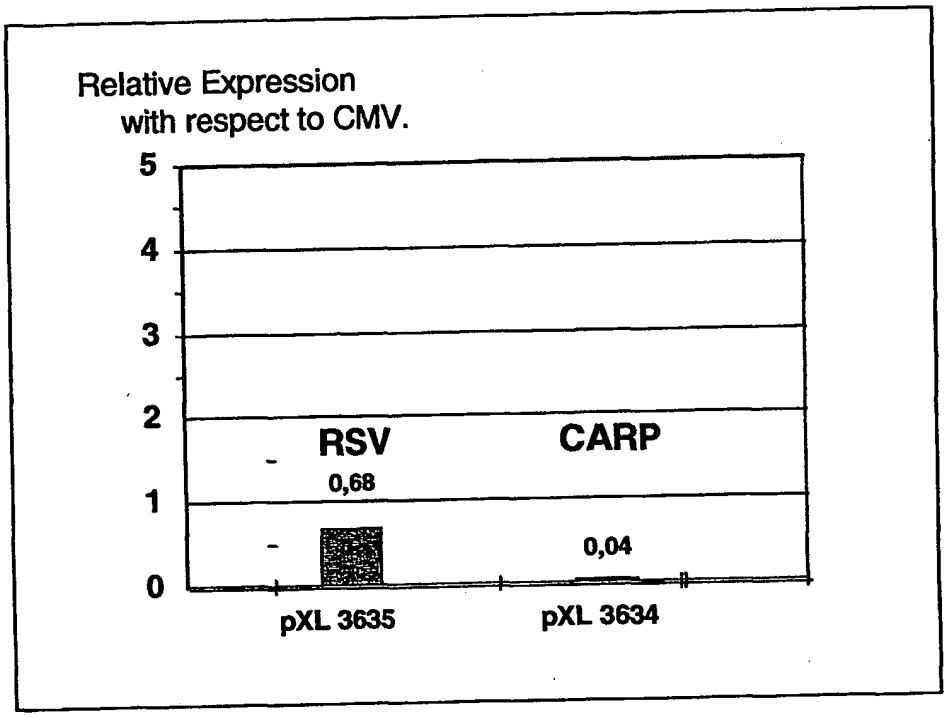

Relative Expression
with respect to CMV.
5
4
3
2
1
0
RSV
CARP
0,68
0,04
pXL 3635
pXL 3634

FIGURE 6A

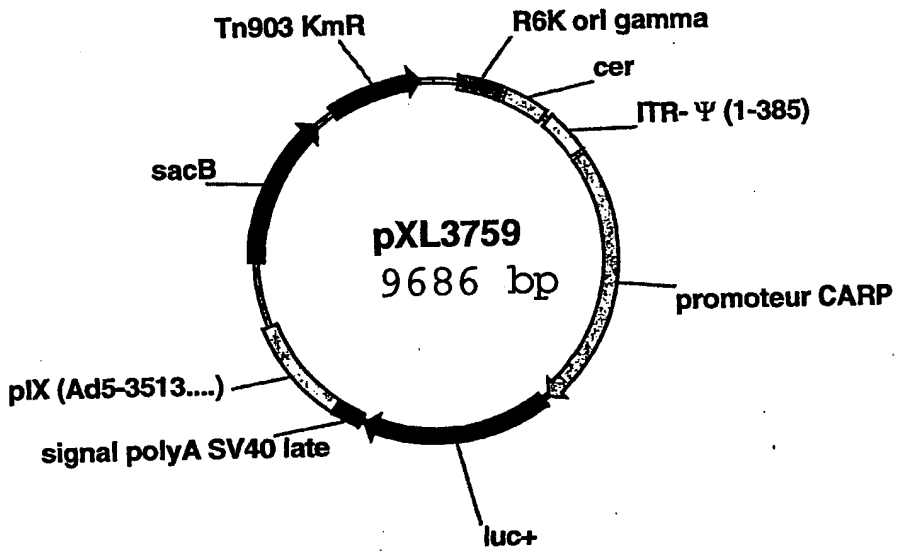


FIGURE 6B

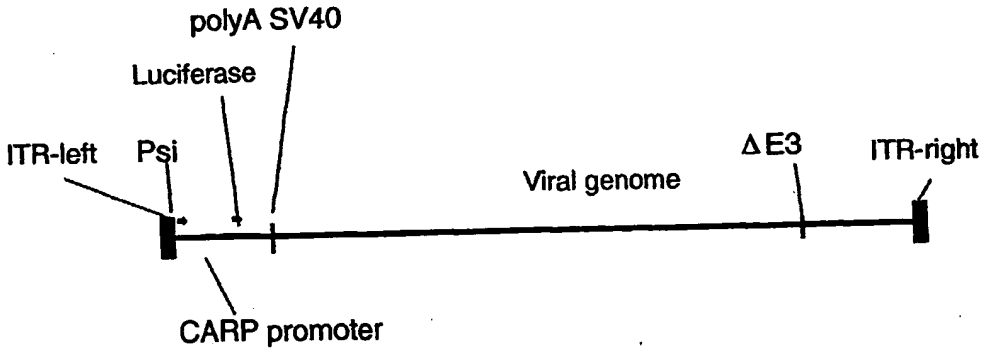


FIGURE 7A

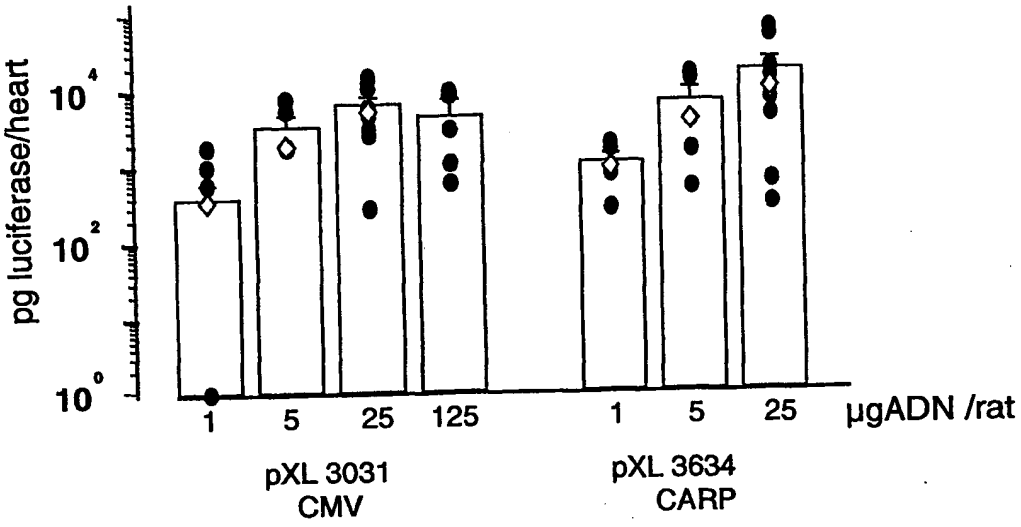


FIGURE 7B

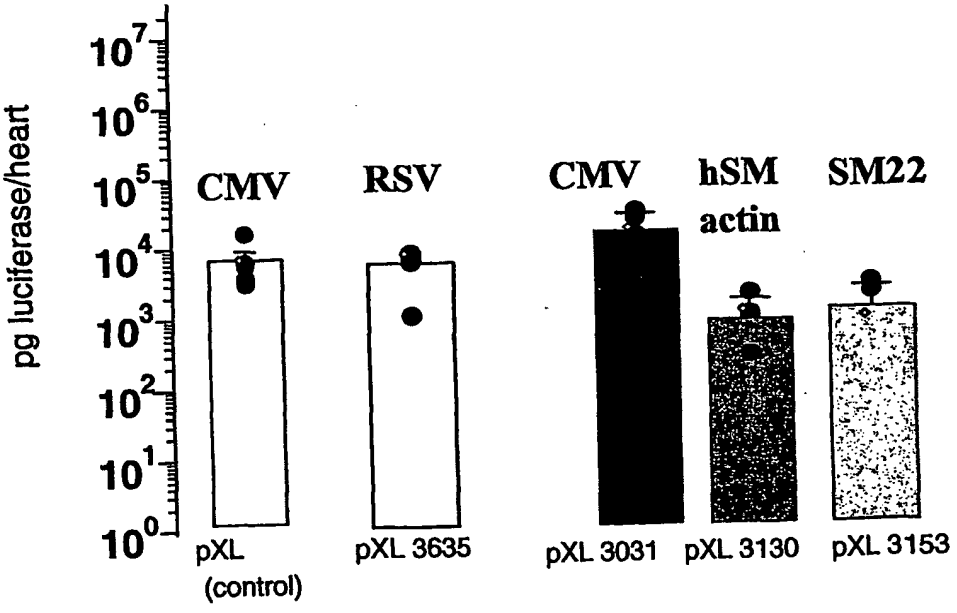
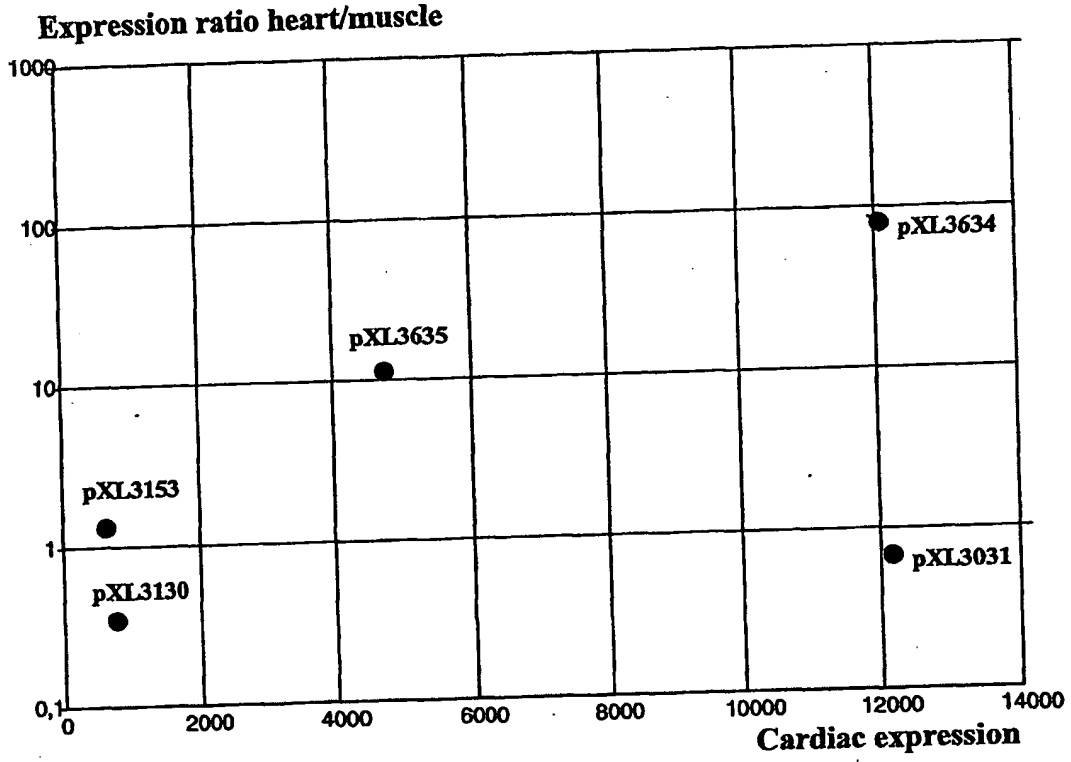


FIGURE 8


Expression ratio heart/muscle
1000
100
10
1
0,1
0
2000
4000
6000
8000
10000
12000
14000
Cardiac expression
pXL3634
pXL3635
pXL3153
pXL3130
pXL3031

SEQUENCES LISTING

<110> AVENTIS PHARMA SA
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

<120> SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING THEM AND USES THEREOF

<130> CART

<140> ST00033
<141> 2000-12-05

<160> 3

<170> PatentIn Ver. 2.1

<210> 1
<211> 2358
<212> ADN
<213> Mus musculus

<400> 1
ggatcctttc atgtttaaca atatcaaccc taacccaagg ggaacagcct gcctgacagt 60
ggctttgcca cccatgaata cttcctagtc tagtccgttt gtgaaactca gcccatccca 120
acacttctgc aagccccatc ctctacaagg tgctcattgg gaatttcctg gagcttctct 180
ttcaggatca gcctgattct agggcagcag ttctcaacct gggggcctcg acccctttgg 240
gggaatcaaa cgacccttta caggggtcac atatcatcta tcctatatgt caggtattta 300
cattacgatt cgtaacagta gcaaaattac aggtatgaaa tagcaatgaa ataattttat 360
gattgaaggt caccacaaca tgaggccgcc acactgttct agagaaaaat cacctgggtg 420
gggaaaggtt tgggaaagcc tttctgtcca ttcttcattc ttcaaagtga tgtgttcaca 480
gaaagccttt cagctgttct gctggggctc ttagtaagtc tgagtaggaa ctgtatgtac 540
caggtctgct tcttatgggt ggagccaaga cgcatcgtgg gtggagcgaa gacgcaacct 600
caccttctag ctctgcatcc atagcaagta gcctaatgtt tctgtgtcta ggtgtcatct 660
ctgtgaatcg agatccttgg ccttgcttga attagggagg cacaaaatac tcagagattc 720
aagactgctc agcagcccag agtccttcct caaaggaaag gtctcaactc tcagcccccc 780
ttagctctga gtcaggcctg gaacaaacgg ccacaggaat gagaaaagct gccatagctg 840
cttgtcactt caagaggtca aagaaaatag tgttaaccat gaaaacgaga agaccaacag 900
ttatccattg atagcgtctc aggacagata ggacagagag aacactagga gaggggaacc 960
cacgaaggac aaggtattag tgtgttggtt ttcagggcaa tgtcttgtac tgaagattct 1020
agaaacacaa tttgctggtt gaacagctga agtggggtgg gggttcttac cccatgttca 1080
tggaagggtg agtgaggaga gacagatata tgatggccag cataacaaac atacacaaca 1140
ccctaattaa cacttccctc ttctactgac acccccttca ctctcctctt tcataaaaaa 1200
taaaaaaagt attttatgtg gctcttacga tagaatcttt cctcgaacta taaaaagatc 1260
taaatattta tatttttcac attttaatat cttagcgatg acaagccaga aacaagtatt 1320
ttttgcctct ctcaacagca aagcttgggg ccttttttgtt tccgtgttag gaatagaaca 1380
cgagagcccc gtgtatctag gcagatgctc tatcattagc ccatgagtct ccagcctcag 1440
acgcacattt ttctcgggct ctcttaagct tttcccacag cattgggaaa ctttactgac 1500
agcatccaag ttgtgcttct gctaagaact ggactcacat ctctctgtgc atcacttcgg 1560
cccgttttgg ggtagatcct ctgattagcc ttcagattta gaacacggtg agcctgtggt 1620
gcactaatta tggccagtga caccatagag tcaaagtgca ttactgaatg ctttcaattt 1680
ctcctaatgc tggtacgatg gcatgtcaca gggccatttt agctgcagac atcactccag 1740
agaattccaa acagatagag acaagtggca cccagaccca tctccttccc ctcgggctga 1800
ttatccccag aaataggatg tcccaaagca acacttccca gccaactgga gtgctgataa 1860
gtccagttat cagaaagata tggctgtaag tgtgatgcac agtgcttgca ttttcttgat 1920
acgttagtca tatgagagct gacaaagaag gaaaaagagc agcgatgtgg tgcaatatta 1980
acaggcagct gtcccctggc ttcccgatac gtgggatgac tcgcattgct gagcggtgtg 2040
gtcactgcca aaggaatgac cctctcacat ttcttcctga ttcgcatacg ccgcggccag 2100

cttgtcatct ccctcttggg cttcccagac actaagtctg gaatgaaaat tcacctgcct 2160
ctgaattggc cactggtggg ggcaggggtg tgacttggct tcccaggctg gaagattatc 2220
tcacccagcc ctagctatat aacgggctgg tgtggagggg ctccacaggg ccagttccag 2280
gggttcatcc acaagagaga aaaacataga ctcgaggtct agggagcttg catgcctgca 2340
ggtcggaggc caccatgg 2358

<210> 2
<211> 2074
<212> ADN
<213> Homo sapiens

<400> 2
ctgcagcaag ttacttaatg ttttttgcct cagcatcctc tctgtaaaat gagagcatta 60
gtcttgctcc aacttcgagg gcatggacag ctctgggatt tcatatccaa gacccttaaa 120
catcccacag tccttccccc aaacacttct cctcctaata cctccctcag tttgggtcag 180
gcctggaaca aaaaggcata cgaaatggta gaaaaagtgt ccatgactac ttctgactta 240
gatgaagaga ccaatgaaaa tagtaatgac tctgtttgct tcagcaggac atatactaaa 300
ataggagcta tacaaagaag attagcatgg actctgtgca agaatgacac acaaatttgt 360
gaaacattcc atatattaaa aataaataaa taataaagag aaaaggaaaa aattaaaaag 420
aaaatagtga tagctgtgtc catctcaaag aaaagcccag gagatttcct ttatttaccc 480
cctttaagat agaatattag gagaccggaa catatgatac aggaggtact gggagggtcc 540
ctctttgtca atgttttgtc ttggggtggg gagtcgatgt cttctcaaag tttcagaaac 600
accatccact gactgagcat tcaaggggca agaggagaat ggcagccaca tttgttgatt 660
gggtgagttt ggggagaaat agacacacaa aggtcaaaca taacttccta attaacactt 720
ccctccattc acaattccct tctcccattc ttctctcctg tcttttacts akaraaaccc 780
agtttttcct gaaactataa aaatacccca agtatgttta cataatttac acctcaaaga 840
ttagaaacca gaaatagaga ccttttcaac ccttccggaa gcaaagtgca ttatccctcc 900
agccacgtgt ctcaaatctt gatgcatcag aatcatctgg gtgctttkaa attcaagatg 960
attcctacga gttaccataa atcaactcag aattccctgg agtggggcca gggatctgta 1020
tttctgacaa gctcccacag gtgattcctt tccccacagc atttgagaac ttcagctcaa 1080
tgacctaatc agagtcctgc cattgctaat atctggtctc atttttbtca tatatatata 1140
tagtatttgt ggtagagatg ggattttgcc atgttgccca ggctagtatt gaactcctaa 1200
gctaagcaat cttcctgtct ctgcctccca aaatgttggg attacaggtg taagccactg 1260
cacccggctg atagctggtt tcatttactc tatttcttga ccactctgat ccattttgaa 1320
gtaaaaatgc tccaattatt atgctgtttt agaacacggt aagcatgtca tgtgctaatg 1380
gccagtgaca tcataaaaga aaagtgcatt actgaatgct ttcaatgtct tataatgatg 1440
gtaaggtggc atgtcatggg gcctatttag cccagacatc actccaaaga attccaaaca 1500
gatatagaca agtgccttta gggcccagat cccttccccc caggctgttt acccagggaa 1560
taggatgtcc tgggacaagt ttcccctaag tgaagtgttg ataagtctgc ttatcagaaa 1620
gatattactg ggggtgtgat atgtagggca tctacatttt cttgataggt agtcatatga 1680
aagctgacaa agaaaaaaag ggcagtgatg tggtgcaatg tcaacagaca gctgtcccct 1740
gactcttgac aaataggatg acttgcattg ctgagcgatg tgatcaccac caaaggaatg 1800
gccctctcac atttcttcct gattcacata ttcagcaggg ttagcttgtc ctcccctccc 1860
tcttcagctt cccagacact gagtctggaa tgaaaattca cctgcctctg agttggctcc 1920
taatgggggc gggagtgtta cttcggttcc caggttggaa gattatctca cccggcccca 1980
gctatataag ctgaccggtg tggaggggcc cagcagggcc aactccaggg attccttcca 2040
cgacagaaaa acatacaaga ctccttcagc caac 2074

<210> 3
<211> 750
<212> PRT
<213> Homo sapiens

<400> 3
Met Gly Glu Thr Leu Gly Asp Ser Pro Ile Asp Pro Glu Ser Asp Ser
1 5 10 15

```
Phe Thr Asp Thr Leu Ser Ala Asn Ile Ser Gln Glu Met Thr Met Val
            20                  25                  30

Asp Thr Glu Met Pro Phe Trp Pro Thr Asn Phe Gly Ile Ser Ser Val
        35                  40                  45

Asp Leu Ser Val Met Glu Asp His Ser His Ser Phe Asp Ile Lys Pro
    50                  55                  60

Phe Thr Thr Val Asp Phe Ser Ser Ile Ser Thr Pro His Tyr Glu Asp
65                  70                  75                  80

Ile Pro Phe Thr Arg Thr Asp Pro Val Val Ala Asp Tyr Lys Tyr Asp
                85                  90                  95

Leu Lys Leu Gln Glu Tyr Gln Ser Ala Ile Lys Val Glu Pro Ala Ser
            100                 105                 110

Pro Pro Tyr Tyr Ser Glu Lys Thr Gln Leu Tyr Asn Lys Pro His Glu
        115                 120                 125

Glu Pro Ser Asn Ser Leu Met Ala Ile Glu Cys Arg Val Cys Gly Asp
    130                 135                 140

Lys Ala Ser Gly Phe His Tyr Gly Val His Ala Cys Glu Gly Cys Lys
145                 150                 155                 160

Gly Phe Phe Arg Arg Thr Ile Arg Leu Lys Leu Ile Tyr Asp Arg Cys
                165                 170                 175

Asp Leu Asn Cys Arg Ile His Lys Lys Ser Arg Asn Lys Cys Gln Tyr
            180                 185                 190

Cys Arg Phe Gln Lys Cys Leu Ala Val Gly Met Ser His Asn Ala Ile
        195                 200                 205

Arg Phe Gly Arg Met Pro Gln Ala Glu Lys Glu Lys Leu Leu Ala Glu
    210                 215                 220

Ile Ser Ser Asp Ile Asp Gln Leu Asn Pro Glu Ser Ala Asp Leu Arg
225                 230                 235                 240

Ala Leu Ala Lys His Leu Tyr Asp Ser Tyr Ile Lys Ser Phe Pro Leu
                245                 250                 255

Thr Lys Ala Lys Ala Arg Ala Ile Leu Thr Gly Lys Thr Thr Asp Lys
            260                 265                 270

Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu Met Met Gly Glu Asp
        275                 280                 285

Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln Glu Gln Ser Lys Glu
    290                 295                 300

Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe Arg Ser Val Glu Ala
305                 310                 315                 320
```

```
Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile Pro Gly Phe Val Asn
                325                 330                 335

Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys Tyr Gly Val His Glu
            340                 345                 350

Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn Lys Asp Gly Val Leu
        355                 360                 365

Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu Phe Leu Lys Ser Leu
    370                 375                 380

Arg Lys Pro Phe Gly Asp Phe Met Glu Pro Lys Phe Glu Phe Ala Val
385                 390                 395                 400

Lys Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp Leu Ala Ile Phe Ile
                405                 410                 415

Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly Leu Leu Asn Val Lys
            420                 425                 430

Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln Ala Leu Glu Leu Gln
        435                 440                 445

Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu Phe Ala Lys Leu Leu
    450                 455                 460

Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr Glu His Val Gln Leu
465                 470                 475                 480

Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met Ser Leu His Pro Leu
                485                 490                 495

Leu Gln Glu Ile Tyr Lys Asp Leu Tyr Ala Trp Ala Ile Leu Thr Gly
            500                 505                 510

Lys Thr Thr Asp Lys Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu
        515                 520                 525

Met Met Gly Glu Asp Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln
    530                 535                 540

Glu Gln Ser Lys Glu Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe
545                 550                 555                 560

Arg Ser Val Glu Ala Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile
                565                 570                 575

Pro Gly Phe Val Asn Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys
            580                 585                 590

Tyr Gly Val His Glu Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn
        595                 600                 605

Lys Asp Gly Val Leu Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu
    610                 615                 620

Phe Leu Lys Ser Leu Arg Lys Pro Phe Gly Asp Phe Met Glu Pro Lys
```

625 630 635 640

Phe Glu Phe Ala Val Lys Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp
645 650 655

Leu Ala Ile Phe Ile Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly
660 665 670

Leu Leu Asn Val Lys Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln
675 680 685

Ala Leu Glu Leu Gln Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu
690 695 700

Phe Ala Lys Leu Leu Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr
705 710 715 720

Glu His Val Gln Leu Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met
725 730 735

Ser Leu His Pro Leu Leu Gln Glu Ile Tyr Lys Asp Leu Tyr
740 745 750